Contents lists available at ScienceDirect

Ecological Indicators

Original article

Macroinvertebrate-based multimetric predictive models for evaluating the human impact on biotic condition of Bolivian streams

Nabor Moya [a,*], Robert M. Hughes [b], Eduardo Domínguez [c], François-Marie Gibon [d], Edgar Goitia [e], Thierry Oberdorff [d,**]

[a] UMR BOREA, IRD 207/CNRS 7208/MNHN/UPMC, Unidad de Limnología y Recursos Acuáticos (ULRA), Universidad Mayor de San Simón, Calle Sucre y Parque la Torre, Casilla 2352, Cochabamba, Bolivia
[b] Department of Fisheries and Wildlife, Oregon State University & Amnis Opes Institute, Corvallis, Oregon 97331, USA
[c] CONICET—Instituto de Biodiversidad Neotropical (IBN), Facultad de Ciencias Naturales, Universidad Nacional de Tucumán, Miguel Lillo 205, 4 000, Tucumán, Argentina
[d] UMR BOREA, IRD 207/CNRS 7208/MNHN/UPMC, DMPA, Muséum National d'Histoire Naturelle, CP 26, # 75231, Paris cedex 05, France
[e] Unidad de Limnología y Recursos Acuáticos (ULRA), Universidad Mayor de San Simón, Cochabamba, Bolivia

ARTICLE INFO

ABSTRACT

We developed and validated a single multimetric index based on predictive models that could evaluate anthropogenic disturbances in streams of three disparate ecoregions of Bolivia. To do so, we examined 45 candidate metrics reflecting different aspects of macroinvertebrate assemblage structure and function gleaned from available literature and for their potential to indicate degradation. More importantly, we integrated functional trait metrics to improve the sensitivity of our index. To quantify possible deviation from reference conditions, we first established and validated statistical models describing metric responses to natural environmental differences in the absence of any significant anthropogenic disturbance. We considered that the residual distributions of these models described the response range of each metric, independently of natural environmental influence. After testing the sensitivity of these residuals to a gradient of anthropogenic disturbance, we retained eight metrics that were used in the final assemblage index, four metrics based on richness and composition and four metrics based on biological traits. Our index performed well in discriminating between reference and disturbed sites, giving a significant negative linear response to a gradient of physical and chemical anthropogenic disturbances. After employing a probability survey design and sampling a relatively small number of sites throughout all major ecoregions of Bolivia, we believe our methodology can be used to develop a monitoring tool to evaluate status and trends in biological condition for streams of the entire country despite its complex and heterogeneous geology and climate.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Rivers and streams are the most intensively human-influenced ecosystems on Earth, partly because they drain effluents from entire catchments and because human population densities and associated activities are highest along river courses (Dudgeon, 2006). As a result, many rivers are severely altered, suffering from channel and bank modifications, flow regulation and fragmentation, chemical pollution and organic pollution (Dudgeon, 2006).

To monitor and manage these ecosystems, it is necessary to develop practical tools using biologically based approaches. These tools should be consistently applicable to different basins and/or ecoregions (Hughes and Peck, 2008). The index of biological integrity (IBI), first introduced by Karr (1981) to assess the health of streams in the US Midwest using fish assemblages, is now applied almost worldwide and for multiple assemblages (Bryce et al., 2002; Hill et al., 2000; Hughes and Oberdorff, 1999; Klemm et al., 2003). The IBI employs a series of metrics based on assemblage structure and function, including metrics directly related to species richness. The rationale behind such a multimetric index is that the different metrics offer robust and sensitive insights into how an assemblage responds to anthropogenic stressors (Karr and Chu, 2000). However, a major difficulty in identifying metrics that best respond to anthropogenic pressures is that these metrics also display natural spatial patterns (Whittier et al., 2006; Ibañez et al., 2009). This has led scientists to adapt specific indices to suit natural regional and local differences, hindering making comparable assessments at large spatial scales (e.g. at regional, national or continental scales; Paulsen et al., 2008; Stoddard et al., 2008; Waite et al., 2010).

* Corresponding author.
** Corresponding author at: IRD, DRV, DMPA, MNHN, 43 rue Cuvier, 75005 Paris, France. Tel.: +33 140793767; fax: +33 140793771.
E-mail addresses: nabor.moya@gmail.com (N. Moya), oberdorf@mnhn.fr (T. Oberdorff).

This methodological limit has been corrected by developing multimetric predictive models of widespread applicability that account for the many possible sources of inter and/or intra-regional variations in assemblage structure that result from variations in natural conditions (e.g. Oberdorff et al., 2001, 2002; Pont et al., 2006, 2009; Tejerina-Garro et al., 2005, 2006; Waite et al., 2010). These multimetric predictive models enable site-specific predictions of metric values, independent of natural environmental factors, and expected in the absence of anthropogenic impact.

In South America, deterioration of aquatic ecosystems is increasing because of recent increases in several economic activities, such as mining and extensive agriculture (Domínguez and Fernández, 2009). For this reason, there is increased interest in using macroinvertebrate assemblages to assess stream biological condition (e.g. Jacobsen, 1998; Barbosa et al., 2001; Marques and Barbosa, 2001; Baptista et al., 2007; Moya et al., 2007; Tomanova et al., 2008; Hepp and Santos, 2009; Mesa, 2010; Miserendino and Masi, 2010). However, the only South American index based on predictive models, and using macroinvertebrate assemblages, was developed by Moya et al. (2007) for some Bolivian streams belonging to a single basin and ecoregion. As such, this index is not applicable elsewhere. Water pollution in Bolivian streams is still considered minimal, but some stream reaches have experienced profound modifications, mostly resulting from deforestation, urbanization, and mining. However, continued population and economic growth in Bolivia will likely increase water pollution, leading to a general deterioration of streams and loss of biological richness (Miller Reed and Czech, 2005; Rose, 2005). Biological indicators that can predict these significant social and ecological impacts are increasingly necessary for minimizing or avoiding future environmental degradation of aquatic resources and the environmental services and human cultures that depend on them.

Therefore, our objective was to develop and validate a single multimetric index applicable to three contrasting ecoregions of Bolivia. If such an index can efficiently discriminate between natural and anthropogenic disturbances, the methodology can be later used in a national probability survey to develop a monitoring tool to evaluate status and trends in the biological condition of streams of the entire country (and eventually neighboring countries).

2. Methods

2.1. Study area

The study area (Fig. 1) covers three important Bolivian ecoregions: the Andean Piedmont, which includes the upper Isiboro-Secure Basin of the Bolivian Amazon, with altitudes of 221–260 m; the Inter-Andean Valleys, which includes the Río Grande Basin, with altitudes of 1505–4286 m; and the High-Andes ecoregion, with altitudes of 3719–4449 m. Across these ecoregions, the climate varies from very humid (5000 mm mean annual precipitation) and warm (26 °C mean annual temperature) to dry (287 mm mean annual precipitation) and cold (7.2 °C mean annual temperature) (Navarro and Maldonado, 2002).

2.2. Site selection

We selected 199 sites within first to fourth-order streams (mean stream width ranging from 1.2 to 32.7 m) fairly evenly distributed across each ecoregion (56 in the Andean Piedmont, 77 in the Inter-Andean Valleys and 66 in the High-Andes ecoregions). To
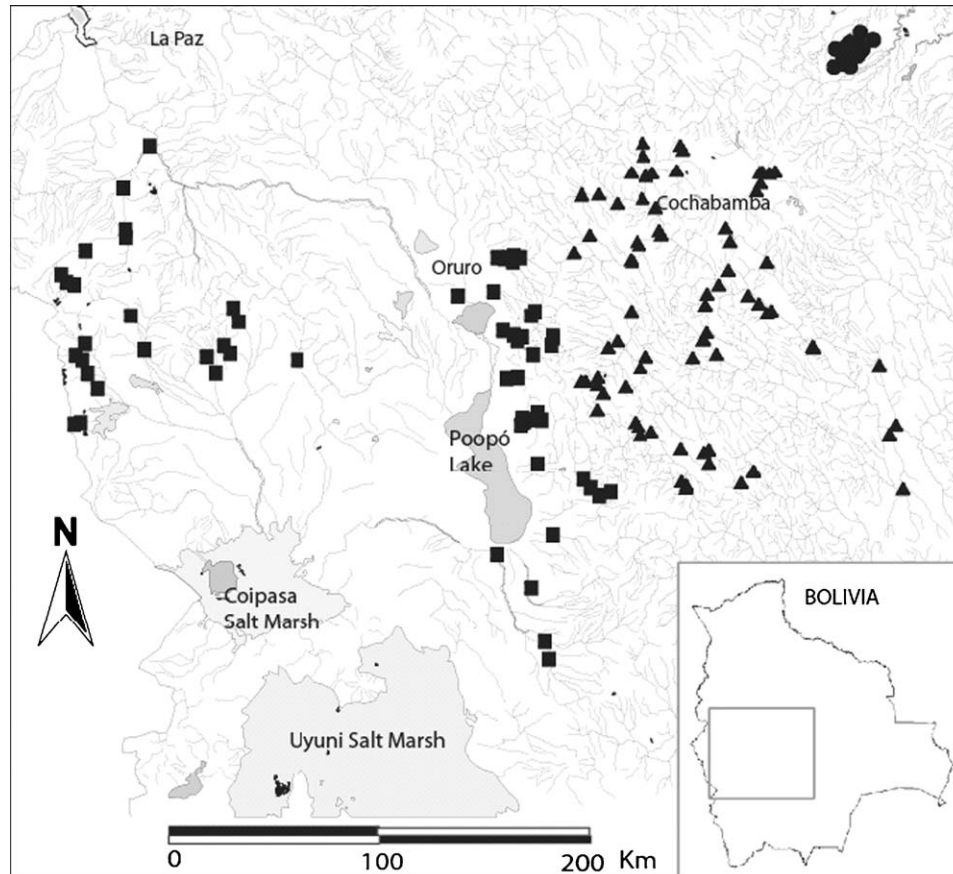
Fig. 1. Location of the 199 sites sampled in the three Bolivian ecoregions (squares = High Andes, triangles = Inter-Andean Valley, and circles = Andean Piedmont).

differentiate between reference and disturbed sites, we first quantified environmental conditions of each site independently of the biota (following the methodology described by Moya et al., 2007). Sites were scored 1, 2, 3, or 4 (from 1, pristine, to 4, highly disturbed) in each of three stressor categories: local habitat modifications (i.e. riparian quality, presence of waste effluents, presence of weirs), landscape use intensity (i.e. urban, deforestation and/or agriculture obtained from land use maps) and mining activities (presence of effluents from mining activities). The three scores were then summed for each site to obtain a final environmental condition score (final scores ranged from 3 to 12). Sites were considered as reference sites if their final environmental condition score was <6, and those sites with scores ≥6 were considered disturbed sites. By using this criterion we artificially include in the reference data set some slightly disturbed sites (i.e. sites with values 4 and 5). However, we chose this procedure to insure a sufficient sample size and geographic distribution to build the metric predictive models. On the basis of these criteria the 199 sites sampled were first divided into two data sets, one data set of 125 reference sites (RS125) and another data set of 74 disturbed sites (DS74). The reference (RS125) data set was further randomly divided into two subsets: one set of 100 sites (RS100) that was used to calibrate the models and another set of 25 sites (RS25) that was used to validate the models. Disturbed sites were selected to cover a range of well-identified anthropogenic effluents, including sewage and urban runoff, agriculture, and mining.

Despite the fact that most altered sites usually mixed at least two categories of disturbances (mostly for urban and agricultural disturbances), we distinguished disturbance categories by classifying sites based on their highest score in each category.

2.3. Sample collection and analysis

Because our goal was to assess the influence of human actions, not natural assemblage variation through time, aquatic macroinvertebrates were collected during the dry season from May to October to standardize the seasonal context. A critical decision in making biological assessments of streams is the choice of the sampling period. Sampling is done preferably when natural variability is at a minimum, when sites are accessible (not flooded, to avoid endangering field crews), and when the human disturbances are most easily detected (low or base flows). This choice also maximizes efficiency of the sampling method. For similar reasons most state or national monitoring programs across the world choose low-flow periods for such sampling (Hughes and Peck, 2008).

High-Andes and Inter-Andean Valleys ecoregions were sampled in 2007 and 2008, respectively. The Piedmont ecoregion was sampled during 2004 (Moya et al., 2007). We used a standard sampling area of approximately $0.09\,m^2$ (Surber sampler, 30 cm × 30 cm × 30 cm, length, width and height, with mesh size 250 μm) and at each site we took five samples having similar depth, flow and substrate from a single riffle as proposed by Resh et al. (1995) and Karr and Chu (1999). Macroinvertebrates and associated material were preserved in 4% formaldehyde. In the laboratory, the samples were washed and all macroinvertebrates were sorted and identified to the family level (Appendix A) through use of Domínguez et al. (2006), Fernández and Domínguez (2001), Merritt and Cummins (1996), and Domínguez and Fernández (2009). The use of family level identification can lead to the loss of potentially important ecological information (Tomanova et al., 2008). However, (i) identification keys at genus or species levels are still limited for neotropical macroinvertebrates, and (ii) the identification of individuals to family level in multimetric index may be a satisfactory alternative, considering that identifying individuals to lower taxonomic levels is highly time consuming (Karr and Chu, 1999; Buss and Vitorino, 2010).

Table 1
Candidate metrics based on assemblage richness and structure, and their expected response to human perturbation ('+' increase '−' decrease). '*' = Metric models not included because they were independent of measured environmental variables.

Metrics	Hypothesized response to human disturbance
Taxa richness and composition	−
Total number of taxa	−
Number of EPT taxa	−
Number of EPT individuals	−
% abundance of EPT individuals*	−
Abundance	
Total number of individuals	−
Trophic composition	
% abundance collector–gatherer individuals*	+
% abundance shredder individuals*	−
% abundance scraper individuals*	−
% abundance predator individuals*	−
% abundance collector–filterer individuals*	+/−
Tolerance	
Number of Chironomidae individuals	+
% abundance of Chironomidae individuals*	+

At each of the five sampling points, we measured several instream variables: current velocity (m/s; measured 5 cm above the bottom), stream depth (cm), stream width (m), conductivity (μs/cm), pH, and substrate heterogeneity (i.e. percentage of each class of substratum particle sizes, following the methodology defined by Heino et al. (2003). Values were then averaged to obtain a mean value for each riffle. All these environmental factors are strong descriptors of physical and hydrological conditions at the local scale and can be considered as important abiotic determinants of richness and structure of local macroinvertebrate assemblages (Vinson and Hawkins, 1998).

2.4. Candidate metrics

Initially, we considered 45 candidate metrics selected primarily from previous studies (Barbour et al., 1996; Gibson et al., 1996; US EPA, 1997) and representing the composition, structure, and functional diversity of aquatic macroinvertebrate assemblages (Tables 1 and 2). Regarding functional metrics, species biological traits were assigned according to Merritt and Cummins (1996) and Tomanova et al. (2006, 2008).

2.5. Metric selection and modeling

Metric selection: The methodology used for metric selection and modeling was mostly derived from Moya et al. (2007), Oberdorff et al. (2002) and Pont et al. (2006). From the 45 candidate metrics, we first removed 18 metrics based on biological traits, because those metric values were never normally distributed (Kolgomoroff–Smirnov test, $p < 0.05$), whatever the transformation used. In a second step, using the reference data set (RS100), each of the 27 remaining metrics was regressed against environmental variables by using stepwise (forward) multiple linear regression procedures and the Akaike information criteria (Hastie and Pregibon, 1993) to build the simplest possible model that adequately explained the candidate metric. The square of each explanatory variable was also included to account for potential non-linear relationships. The three ecoregions (i.e. Andean Piedmont, Inter-Andean Valleys and High-Andes) were coded and entered in the models as nominal (categorical) variables. Some data transformations were done prior to analyses to satisfy statistical assumptions (see Table 3). In the third step, using the reference data set (RS100), we verified that the residual metric values were normally distributed. Fourth, to maximize the independence among

Table 2
Biological traits, categories and expected trends with perturbation ('+' increase, '−' decrease); after Tomanova et al. (2008). * = Metric models not included because they were independent of measured environmental variables, † = metrics excluded because they were originally not normally distributed.

Biological trait	Category	Trend
Food	Sediment particles (S)	+
	Fine detritus (FPOM) <1 mm (FPOM)	+
	Coarse detritus (CPOM) > 1 mm* (CPOM)*	−
	Microphytes (MiPh)	−
	Macrophytes (MaPH)*	−
	Dead Animals (DA) †	−
	Microinvertebrates (MIIn) †	−
	Macroinvertebrates (MAIn)*	+
Respiration	Integument	+
	Gill*	−
	Plastron†	−
	Stigmata†	+/−
Maximal body size (mm)	<2.5†	−
	2.5–5†	+/−
	5–10†	+
	10–20†	−
	20–40†	−
	40–80†	+/−
	>80†	+/−
Body flexibility (degrees)	None (<10) †	+/−
	Low (10–45) †	−
	High (>45) †	+
Body form	Streamlined*	−
	Flattened	−
	Cylindrical*	+
	Spherical†	+/−
Mobility and attachment substrate	Fliers†	−
	Surface swimmer (SwS)*	+/−
	Full water swimmer (SwW) †	−
	Crawler (CL) †	−
	Epibenthic burrower (EpB)	+
	Endobenthic burrower (EnB)*	+
	Temporarily attached (TA)*	+/−

metrics in the final index, we checked for redundancy (metrics with a Pearson's $r > 0.80$ or $<−0.80$ were considered redundant). In cases of metric redundancy we selected the metric showing the strongest response to disturbance (Hughes et al., 1998; Stoddard et al., 2008). Finally, we retained only metrics sensitive to perturbations by verifying that their mean standardized residual values in DS74 were significantly different from zero (t-tests).

Index calculation and validation: A detailed methodology for index development is available in Oberdorff et al. (2002). The residual value of a metric for a given site can be used to test the null hypothesis "the residual value belongs to the reference distribution" against the alternative hypothesis "the residual value differs from the reference distribution in the way expected with

perturbation". To do so, metric residuals were transformed into probabilities according to their response trends with perturbation. All transformed metric values (i.e. residual values transformed into probabilities) vary between zero and one, and decrease as human perturbations increase (Pont et al., 2006). The expected distribution of these probabilities for reference sites is a uniform distribution with a mean value of 0.5. To calculate a site's index score, we added probabilities of each of the retained metrics and divided by the number of metrics to produce an index with scores of 0–1. The index was validated using two independent data sets: a data set of reference sites (RS25) and a data set of disturbed sites (DS74). If our approach is valid, the mean value of the index in the calibration data set (RS100) (i.e. an expected mean index value of 0.5) should not differ statistically from the mean index value of the validation data set (RS25), whereas for the perturbation data set (DS74), the mean value should be significantly lower than those of the calibration and validation data sets (unilateral t-test). All statistical analyses were performed using SYSTAT 12.0 software.

3. Results

Our metric screening process efficiently reduced the set of candidate metrics. From the initial 45 metrics, 18 metrics were excluded because they were never normally distributed, whatever the transformation used (Kolgomoroff–Smirnov test, $p < 0.05$) (Table 2). From the remaining 27 metrics included in the multiple linear regressions, 16 metrics were excluded because none of the environmental descriptors was retained in the models (see Tables 1 and 2 where excluded metrics are marked with an asterisk). Two of the remaining 11 metrics, percent abundance of epibenthic burrowers and percent abundance of sediment particle feeders were highly correlated with each other ($r = 0.804$). We excluded the relative abundance of epibenthic burrowers metric because it showed a weaker response to perturbations. Two of the remaining 10 metrics were removed because they were statistically unresponsive to perturbation (mean of the standardized residual values applied to DS74 not differing from zero), number of Chironomidae ($t = −2.243$, $p = 0.056$) and percent abundance of individuals with tegument respiration ($t = 0.330$, $p = 0.146$).

We finally retained eight metrics, four metrics based on assemblage richness and abundance and four metrics based on assemblage biological traits (Table 3). All metrics were predicted by at least two natural environmental variables, with ecoregion and conductivity being the most common predictors. These eight metrics were not strongly correlated (highest correlation coefficient, $r = 0.64$).

The total score of the final index was obtained by summing combined probabilities corresponding to the eight remaining metrics and divided by eight (number of metrics) to produce index scores from 0 to 1. As expected, the mean of the index scores for the calibration data set did not differ from the expected value of 0.5

Table 3
Stepwise multiple regression models obtained for the eight metrics included in the final index. FPOM = fine detritus <1 mm. p values<0.01 in all models. Metric and environmental variable transformations to satisfy the assumptions of statistical analyses are given in parentheses.

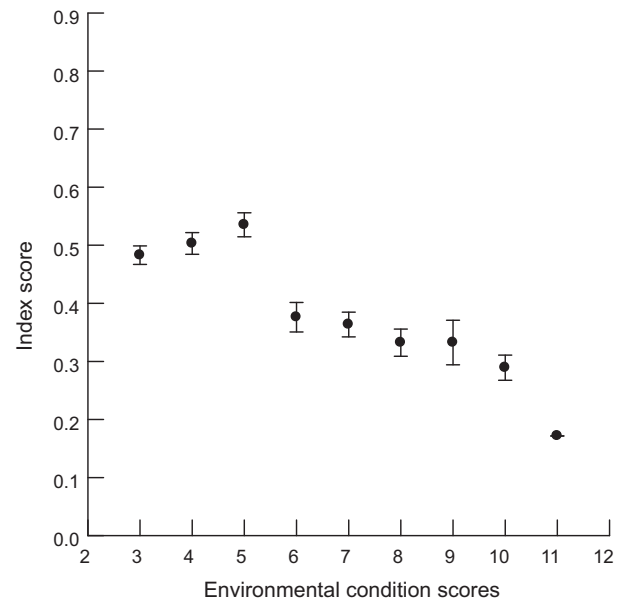| Metric	Total abundance (ln + 1)	Total richness	EPT richness (ln + 1)	EPT abundance (ln + 1)	% sediment	% microphyte	% FPOM	% Flattened ($\sqrt{}$)
Intercept	10.017	20.982	1.667	7.943	14.351	10.1	50.289	3.748
Piedmont region	1.083	−2.152	−0.345				1.933	
Valley region				−0.393	−1.368			0.693
pH	−0.378							
Conductivity (ln + 1)		−1.186			−0.653	1.083	−1.25	
Mean velocity (ln + 1)						5.44		−2.854
Mean width (ln + 1)			0.121					
Mean depth (ln + 1)				−0.778				
Substrate			0.083					
r^2	0.39	0.27	0.61	0.16	0.19	0.14	0.16	0.37
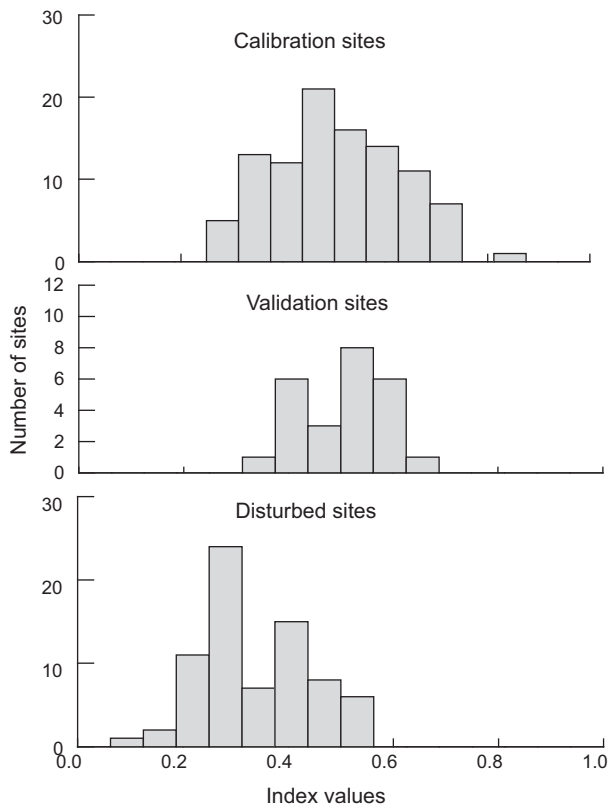
Fig. 2. Distribution of the index values for the calibration sites [RS100], the validation sites [RS25] and the disturbed sites [DS64].

Fig. 3. Relationships between mean index scores (±SE) and environmental condition scores (ECS) for the calibration sites [RS100], the validation sites [RS25] and the disturbed sites [DS64]. The ECS for reference sites are <6, and for disturbed sites >6.

(0.505, $p = 0.416$) and was not significantly different than that of the validation data set (0.504) ($t = 0.067$, $p = 0.947$), whereas the mean of the index scores for the disturbed data set (0.339) was significantly lower than that of the calibration ($t = -13.707$, $p < 0.0001$) and validation ($t = -10.157$, $p < 0.0001$) data sets, indicating clear discrimination between reference and impaired sites (Fig. 2). A highly significant negative relationship was observed between mean index values (overall data set) and independent environmental condition scores (ECS) ($n = 199$, $r = -0.50$, $p < 0.0001$), even if environmental scores 6–7 and 8–9 could not be statistically discriminated (Fig. 3). Furthermore, despite the fact that mean index scores were overall invariant for all unimpaired sites, there was a slightly higher mean index score at ECS = 5 (0.53), compared to ECS = 3 (0.498) and ECS = 4 (0.50) (Fig. 3). This may be the result of an enrichment effect whereby a slight increase in nutrients or temperature in a cool oligotrophic system produces more species and individuals (Davies and Jackson, 2006). However, this did not affect the results, because the overall mean index score value for all unimpaired sites did not statistically differ from the expected value of 0.5 (0.505, $p = 0.318$). A one-way ANOVA used to test for differences in final index values for the three disturbance categories (i.e. agriculture, urban, and mining) indicated that index values were significantly different among those categories ($F = 6.884$, $n = 74$, $p = 0.002$), with an overall decrease in index values from agricultural to mining disturbances (Fig. 4).

4. Discussion

We developed an ecological index that accounts for the major possible sources of anthropogenic and natural inter- and intra-regional variation in macroinvertebrate assemblage structure between the High-Andes and Andean Piedmont of Bolivia. Our index combines several types of metrics related to richness, struc-

ture, and function of these assemblages. The use of functional trait metrics should improve the robustness of our index, because such traits are sensitive to anthropogenic disturbances and are comparable among assemblages even across ecoregions that differ in their taxonomic composition (e.g. Archaimbault et al., 2010; Dolédec and Statzner, 2010; Lamouroux et al., 2004; Statzner and Bêche, 2010; Statzner et al., 2001; Tomanova et al., 2008).

The final index includes only 8 metrics out of the 45 original candidates. Like Klemm et al. (2003), Pont et al. (2006) and Stoddard et al. (2008), we drastically reduced the number of metrics, eliminating those that were highly variable or unresponsive. We screened metrics by testing the performance of each of the 45 candidate metrics with two independent data sets of reference and disturbed sites, finally retaining only the most effective ones (Table 3). For example, many of the metrics based on biological traits were excluded because they were not normally distributed. We also excluded all the feeding group metrics (i.e. % collector–gatherer individuals, % shredder individuals, % scraper
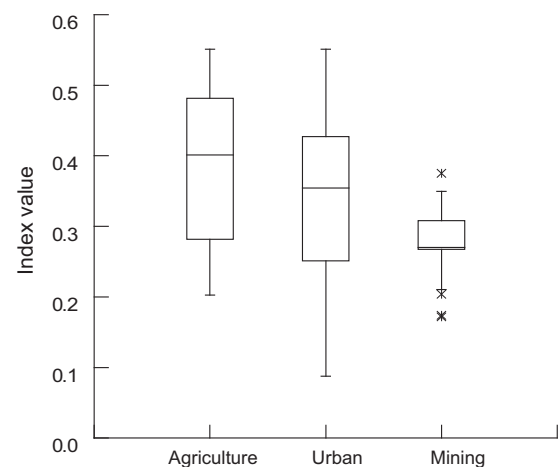
Fig. 4. Index values for the three disturbance types (ANOVA test, $F = 6.884$, $n = 74$, $p = 0.002$). Box represents the 25th and 75th percentiles, horizontal line within a box represents the median, whiskers represent the 5th and 95th percentiles.

individuals, % predator individuals, % collector–filterer individuals; Table 1) because they did not respond to perturbations, contrary to some of their related functional trait metrics (i.e. % sediment particles feeders, % fine detritus feeders, % microphyte feeders). Two main explanations can be advanced to explain this apparent dichotomy. First, there is considerable within-family variation in the feeding habits of macroinvertebrate species in neotropical streams (Moya et al., 2007; Palmer et al., 1996; Thorne and Williams, 1997) and the use of family level taxonomy (as in our study) can thus lead to a great loss of ecological information. Secondly, biological traits are reported to be less sensitive than feeding groups to taxonomic resolution level (Dolédec et al., 2000; Gayraud et al., 2003).

The two metrics, number and percent of Chironomidae are presumed to increase with domestic and agriculture disturbance (Moya et al., 2007; Tomanova et al., 2008); however, we found that they were not able to discriminate between natural and human disturbance. Those metrics were not useful in xeric or mountain ecoregions of the USA either (Klemm et al., 2003; Whittier et al., 2006). On the other hand, Whittier et al. (2006) reported that in the sand-bottomed streams of the Northern Plains Ecoregion of the USA, % Chironomid individuals scored higher in reference than in disturbed sites. Chironomids are a very diverse taxon including species with extremely different pollution sensitivities; consequently they show high variation in relative abundance according to environmental conditions (Barbour et al., 1992), suggesting that Chironomids must be identified to a taxonomic unit whose pollution tolerance is known. However, because Chironomid taxonomy is time consuming, they may not be cost-effective for use in IBIs (Kerans and Karr, 1994), particularly if other invertebrates are easier to identify and sensitive to anthropogenic disturbances. For our study, a possible explanation for the absence of a Chironomid metric response to perturbation is that we included sites disturbed by mining activities generally producing much stronger effects than other type of disturbances (see Section 3). Particularly in the high-Andean region, we found some sites without any macroinvertebrates as a consequence of acid mine drainage. This means that even tolerant taxa like Chironomids decrease or disappear in sites strongly altered by mining activities (Van Damme et al., 2008).

The eight metrics retained represent four classes of biological attributes. Two metrics, based on species richness and composition (i.e. total richness and EPT richness) measure the extent to which elements of biotic diversity are present. Three metrics based on macroinvertebrate feeding traits (i.e. % sediment particle feeding individuals, % microphyte feeding individuals and % coarse detritus feeding individuals) evaluate assemblage condition relative to food particle size and type. One trait metric (i.e. % of flattened body individuals) assesses benthic habitat conditions (sediment-free coarse substrate; Statzner and Bêche, 2010). Two metrics (i.e. total abundance and EPT abundance) evaluate assemblage size. Consequently, our index incorporates several assemblage characteristics and covers a broad range of anthropogenic impacts.

Our index performed well in discriminating between reference and disturbed sites, showing a stable response for all unimpaired sites whatever the natural environmental condition and a significant negative linear response along a gradient of physical and chemical human disturbances. Moreover, it can be applied in three distinct ecoregions of Bolivia through use of a consistent set of metrics, despite the complex and heterogeneous geology and climate of this country. Consequently, based on this methodology, and by enlarging the spatial extent of sampling sites to all of Bolivia, it will be possible to (i) develop a single multimetric index to monitor change, (ii) establish realistic national, regional and land use specific benchmarks or criteria for index scores, and (iii) provide a baseline for predicting and measuring the full biotic response to future rehabilitation or further degradation of Bolivian streams. The development of such an index offers an opportunity to facilitate unbiased and accurate national stream biological assessments.

Another bioassessment method, RIVPACS (River InVertebrate Prediction And Classification System), is also a multivariate predictive model that aids detection and interpretation of anthropogenic stress on invertebrate assemblage richness of streams and rivers separately from natural gradients (Wright et al., 1984). RIVPACS models are based on empirical relationships between individual taxon probabilities of capture and natural environmental features (e.g. substrate composition, alkalinity, elevation, etc.) that are derived from data collected from a reference site network. The main difference from our approach is that RIVPACS uses only taxonomic richness to assess a site's overall condition instead of several metrics as we propose. Similar to the approach of Oberdorff et al. (2002) and Pont et al. (2006, 2009), our species richness metric may thus be considered a RIVPACS-type descriptor inserted into a more general, multimetric, approach. Because each metric reflects the quality of a different aspect of the macroinvertebrate assemblage that could respond in a different manner to aquatic ecosystem stressors, the combination of metrics (as proposed in our index) should be more comprehensive than a single richness measure for estimating the ecological condition of these complex ecological systems. In other words, a multimetric predictive model combines the strengths of the IBI (Hughes and Oberdorff, 1999; Karr, 1981; Roset et al., 2007) and RIVPACS (Hawkins et al., 2000; Paulsen et al., 2008; Wright et al., 1984) approaches for assessing and predicting assemblage condition at national geographic extents.

Appendix A. List of taxa found in the three Bolivian ecoregions. Superscript numbers '1,2,3' indicate occurrence of taxa for Piedmont, Inter-Andean Valley or High-Andean ecoregions respectively. Taxa without number indicate their occurrence in all three ecoregions.

Taxa			
Baetidae	Sericostomatidae[2]	Gomphidae	Veliidae
Leptophlebiidae	Heteroceridae[2]	Calopterygidae[1]	Saldidae[3]
Leptohyphidae[1,2]	Hydraenidae[2]	Aeshnidae[2,3]	Gerridae[1]
Caenidae[1,2]	Dytiscidae	Chironomidae	Belostomatidae[1,2]
Oligoneuriidae[3]	Gyrinidae[2]	Simuliidae	Pyralidae
Grypopterigidae[2,3]	Chrysomelidae[2]	Tipulidae	Corydalidae[1]

Appendix A (*Continued*)

Taxa			
Perlidae[1,2]	Ptilodactylidae[1,2]	Empididae	Sialidae[1]
Hydropsychidae	Elmidae (larvae)	Muscidae	Hyalellidae[2,3]
Hydroptilidae	Elmidae (adult)	Ceratopogonidae	Ancylidae[2]
Calamoceratidae[1,2]	Psephenidae[1]	Ephydridae[2,3]	Planariidae
Philopotamidae[1,2]	Hydrophilidae (larvae)	Tabanidae	Hirudinea
Xiphocentronidae[1]	Hydrophilidae (adult)	Athericidae[2,3]	Acarina
Glossosomatidae[1,3]	Dytiscidae (larvae)[2,3]	Dolichopodidae	Collembola
Odontoceridae[1]	Dytiscidae (adult)[2,3]	Psychodidae	Bivalvia
Polycentropodidae[1,2]	Staphylinidae[2,3]	Stratiomyidae	Gastropoda
Helicopsychidae	Curculionidae (adult)[3]	Limonidae[3]	Ostracoda
Leptoceridae[1,3]	Coenagrionidae	Naucoridae	Nematoda
Limnephilidae[2,3]	Libellulidae[1]	Corixidae[2,3]	Oligochaeta
Hydrobiosidae[2,3]			

References

Archaimbault, V., Usseglio-Polatera, P., Garric, J., Wasson, J.-G., Babut, M., 2010. Assessing pollution of toxic sediment in streams using bio-ecological traits of benthic macroinvertebrates. Freshwater Biology 55, 1430–1446.

Baptista, D.F., Buss, D.F., Egler, M., Giovanelli, A., Silveira, M.P., Nessimian, J.L., 2007. A multimetric index based on benthic macroinvertebrates for evaluation of Atlantic Forest streams at Rio de Janeiro State, Brazil. Hydrobiologia 575, 83–94.

Barbosa, F.A.R., Callisto, M., Galdean, N., 2001. The diversity of benthic macroinvertebrates as an indicator of water quality and ecosystem health: a case study for Brazil. Aquatic Ecosystem Health and Management 4, 51–59.

Barbour, M.T., Plafkin, J.L., Bradley, B.P., Graves, C.G., Wisseman, R.W., 1992. Evaluation of EPA's rapid bioassessment benthic metrics: metric redundancy and variability among reference stream sites. Environmental Toxicology and Chemistry 11, 437–449.

Barbour, M.T., Gerritsen, J., Griffith, G.E., Frydenborg, R., McCarron, E., White, J.S., Bastian, M.L., 1996. A framework for biological criteria for Florida streams using benthic macroinvertebrates. Journal of the North American Benthological Society 15, 185–211.

Bryce, S.A., Hughes, R.M., Kaufmann, P.R., 2002. Development of a bird integrity index: using birds assemblages as indicators of riparian condition. Environmental Management 30, 294–310.

Buss, D.F., Vitorino, A.S., 2010. Rapid bioassessment protocols using benthic macroinvertebrates in Brazil: evaluation of taxonomic sufficiency. Journal of the North American Benthological Society 29, 562–571.

Davies, S.P., Jackson, S.K., 2006. The biological condition gradient: a descriptive model for interpreting change in aquatic ecosystems. Ecological Applications 16, 1251–1266.

Dolédec, S., Statzner, B., 2010. Responses of freshwater biota to human disturbance: contribution of J-NABS to developments in ecological integrity assessments. Journal of the North American Benthological Society 29, 286–311.

Dolédec, S., Olivier, J.M., Statzner, B., 2000. Accurate description of the abundance of taxa and their biological traits in stream invertebrate communities: effects of taxonomic and spatial resolution. Archiv für Hydrobiologie 148, 25–43.

Domínguez, E., Fernández, H.R., 2009. Macroinvertebrados bentónicos sudamericanos: Sistemática y Biología. Fundación Miguel Lillo, Tucumán, Argentina.

Domínguez, E., Molineri, C., Pescador, M.L., Hubbard, M.D., Nieto, C., 2006. Ephemeroptera of South America. In: Adis, J., Arias, J.R., Rueda-Delgado, G., Wantzen, K.M. (Eds.), Aquatic Biodiversity in Latin America. Pensoft Press, Moscow, pp. 650.

Dudgeon, D., 2006. The impacts of human disturbance on stream benthic invertebrates and their drift in North Sulawesi, Indonesia. Freshwater Biology 51, 1710–1729.

Fernández, H.R., Domínguez, E., 2001. Guía para la Determinación de Artrópodos Sudamericanos. Editorial Universitaria de Tucumán, Tucumán, Argentina.

Gayraud, S., Statzner, B., Bady, P., Haybachp, A., Scholl, F., Usseglio-Polatera, P., Bacchi, M., 2003. Invertebrate traits for the biomonitoring of large European rivers: an initial assessment of alternative metrics. Freshwater Biology 48, 1–20.

Gibson, G.R., Barbour, M.T., Stribling, J.B., Gerritsen, J., Karr, J.R., 1996. Biological Criteria: Technical Guidance for Streams and Small Rivers. EPA 822-B-96-001. US Environmental Protection Agency, Office of Water, Washington, DC.

Hastie, T.J., Pregibon, D., 1993. Generalized linear models. In: Chambers, J.M., Hastie, T.J. (Eds.), Statistical Models in S. Chapman & Hall, London, U.K., pp. 195–247.

Hawkins, C.P., Norris, R.H., Hogue, J.N., Feminella, J.W., 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. Ecological Applications 10, 1456–1477.

Heino, J., Muotka, T., Paavola, R., 2003. Determinants of macroinvertebrate diversity in headwater streams: regional and local influences. Journal of Animal Ecology 72, 425–434.

Hepp, L.U., Santos, S., 2009. Benthic communities of streams related to different land uses in a hydrographic basin in southern Brazil. Environmental Monitoring and Assessment 157, 305–318.

Hill, B.H., Herlihy, A.T., Kaufmann, P.R., Stevenson, R.J., McCormick, F.H., Johnson, C.B., 2000. Use of periphyton assemblage data as an index of biotic integrity. Journal of the North American Benthological Society 19, 50–67.

Hughes, R.M., Oberdorff, T., 1999. Applications of IBI concepts and metrics to waters outside the United States and Canada. In: Simon, T.P. (Ed.), Assessing the Sustainability and Biological Integrity of Water Resources using Fish Communities. CRC Press, Boca Raton, Florida, pp. 79–93.

Hughes, R.M., Peck, D.V., 2008. Acquiring data for large aquatic resource surveys: the art of compromise among science, logistics, and reality. Journal of the North American Benthological Society 27, 837–859.

Hughes, R.M., Kaufmann, P.R., Herlihy, A.T., Kincaid, T.M., Reynolds, L., Larsen, D.P., 1998. A process for developing and evaluating indices of fish assemblage integrity. Canadian Journal of Fisheries and Aquatic Sciences 55, 1618–1631.

Ibañez, C., Belliard, J., Hughes, R.M., Irz, P., Kamdem-Toham, A., Lamouroux, N., Tedesco, P.A., Oberdorff, T., 2009. Convergence of temperate and tropical stream fish assemblages. Ecography 32, 658–670.

Jacobsen, D., 1998. The effect of organic pollution on the macroinvertebrate fauna of Ecuadorian highland streams. Archiv für Hydrobiologie 143, 179–195.

Karr, J.R., 1981. Assessment of biotic integrity using fish communities. Fisheries 6, 21–27.

Karr, J.R., Chu, E.W., 1999. Restoring life in Running Waters: Better Biological Monitoring. Island Press, Washington, DC.

Karr, J.R., Chu, E.W., 2000. Sustaining living rivers. Hydrobiologia 422/423, 1–14.

Kerans, B.L., Karr, J.R., 1994. A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. Ecological Applications 4, 768–785.

Klemm, D.J., Blocksom, K.A., Fulk, F.A., Herlihy, A.T., Hughes, R.M., Kaufmann, P.R., Peck, D.V., Stoddard, J.L., Thoeny, W.T., Griffith, M.B., Davis, W.S., 2003. Development and evaluation of a macroinvertebrate biotic index (MBII) for regionally assessing mid-Atlantic Highlands streams. Environmental Management 31, 656–669.

Lamouroux, N., Dolédec, S., Gayraud, S., 2004. Biological traits of stream macroinvertebrate communities: effects of microhabitat, reach, and basin filters. Journal of the North American Benthological Society 23, 449–466.

Marques, M.M., Barbosa, F., 2001. Biological quality of waters from an impacted tropical watershed (middle Rio Doce basin, southeast Brazil), using benthic macroinvertebrate communities as an indicator. Hydrobiologia 457, 69–76.

Merritt, R., Cummins, K., 1996. Aquatic Insects of North America, second ed. Kendall & Hunt Publishing Company, Dubuque, Iowa.

Mesa, L.M., 2010. Effect of spates and land use on macroinvertebrate community in neotropical Andean streams. Hydrobiologia 641, 85–95.

Miller Reed, K., Czech, B., 2005. Causes of fish endangerment in the United States, or the structure of the American economy. Fisheries 30 (7), 36–38.

Miserendino, M.L., Masi, C.I., 2010. The effect of land use on environmental features and functional organization of macroinvertebrate communities in Patagonian low order streams. Ecological Indicators 10, 311–319.

Moya, N., Tomanova, S., Oberdorff, T., 2007. Initial development of a multi-metric index based on aquatic macroinvertebrates to assess streams condition in the Upper Isiboro-Sécure Basin, Bolivian Amazon. Hydrobiologia 589, 107–116.

Navarro, G., Maldonado, M., 2002. Geografía ecológica de Bolivia: vegetación y ambientes acuáticos. Centro de Ecología Simón I. Patiño-Departamento de difusión. Cochabamba, Bolivia, pp. 77–91.

Oberdorff, T., Pont, D., Hugueny, B., Chessel, D., 2001. A probabilistic model characterizing fish assemblages of French rivers: a framework for environmental assessment. Freshwater Biology 46, 399–415.

Oberdorff, T., Pont, D., Hugueny, B., Porcher, J.P., 2002. Development and validation of a fish-based index (FBI) for the assessment of "river health" in France. Freshwater Biology 47, 1720–1735.

Palmer, C.G., Maart, B., Palmer, A.R., O'Keeffe, J.H., 1996. An assessment of macroinvertebrate functional feeding groups as water quality indicators in the Buffalo River, Eastern Cape Province, South Africa. Hydrobiologia 318, 153–164.

Paulsen, S.G., Mayio, A., Peck, D.V., Stoddard, J.L., Tarquinio, E., Holdsworth, S.M., Van Sickle, J., Yuan, L.L., Hawkins, C.P., Herlihy, A.T., Kaufmann, P.R., Barbour, M.T., Larsen, D.P., Olsen, A.R., 2008. Condition of stream ecosystems in the US: an overview of the first national assessment. Journal of the North American Benthological Society 27, 812–821.

Pont, D., Hugueny, B., Beier, B., Goffaux, D., Melcher, A., Noble, R., Rogers, C., Roset, N., Schmutz, S., 2006. Assessing river biotic condition at the continental scale: a European approach using functional metrics and fish assemblages. Journal of Applied Ecology 43, 70–80.

Pont, D., Hughes, R.M., Whittier, T.R., Schmutz, S., 2009. A predictive index of biotic integrity model for aquatic vertebrate assemblages of Western U.S. streams. Transactions of the American Fisheries Society 138, 292–305.

Resh, V.H., Norris, R.H., Barbour, M.T., 1995. Design and implementation of rapid assessment approaches for water resource monitoring using benthic macroinvertebrates. Australian Journal of Ecology 20, 108–121.

Rose, C.A., 2005. Economic growth as a threat to fish conservation in Canada. Fisheries 30, 36–38.

Roset, N., Grenouillet, G., Goffaux, D., Pont, D., Kestemont, P., 2007. A review of existing fish assemblage indicators and methodologies. Fisheries Management and Ecology 14, 393–405.

Statzner, B., Bêche, L.A., 2010. Can biological invertebrate traits resolve effects of multiple stressors on running water ecosystems? Freshwater Biology 55, 80–119.

Statzner, B., Hildrew, A.G., Resh, V.H., 2001. Species traits and environmental constraints: entomological research and the history of ecological theory. Annual Review of Entomology 46, 291–316.

Stoddard, J.L., Herlihy, A.T., Peck, D.V., Hughes, R.M., Whittier, T.R., Tarquinio, E., 2008. A process for creating multi-metric indices for large-scale aquatic surveys. Journal of the North American Benthological Society 27, 878–891.

Tejerina-Garro, F.L., Maldonado, M., Ibañez, C., Pont, D., Roset, N., Oberdorff, T., 2005. Effects of natural and anthropogenic environmental changes on riverine fish assemblages: a framework for ecological assessment of rivers. Brazilian Archives of Biology and Technology 48, 91–108.

Tejerina-Garro, F.L., de Mérona, B., Oberdorff, T., Hugueny, B., 2006. A fish-based index of large river quality for French Guiana (South America): method and preliminary results. Aquatic Living Resources 19, 31–46.

Thorne, R.St.J., Williams, W.P., 1997. The response of benthic macroinvertebrates to pollution in developing countries: a multimetric system of bioassessment. Freshwater Biology 37, 671–687.

Tomanova, S., Goitia, E., Helezic, J., 2006. Trophic levels and functional feeding groups of macroinvertebrates in neotropical streams. Hydrobiologia 556, 251–264.

Tomanova, S., Moya, N., Oberdorff, T., 2008. Using macroinvertebrate biological traits for assessing biotic integrity of neotropical streams. River Research and Applications 24, 1230–1239.

US EPA, 1997. Monitoring Guidance for Determining the Effectiveness of Nonpoint Source Controls. EPA 841-B-96-004. US. Environmental Protection Agency, Office of Water, Washington, DC.

Van Damme, P.A., Hamel, C., Ayala, A., Bervoets, L., 2008. Macroinvertebrate community response to acid mine drainage in rivers of the High Andes (Bolivia). Environmental Pollution 156, 1061–1068.

Vinson, M.R., Hawkins, C.P., 1998. Biodiversity of stream insects: variation at local, basin and regional scales. Annual Review of Entomology 43, 271–293.

Waite, I.R., Brown, L.R., Kennen, J.G., May, J.T., Cuffney, J.L.O., Jones, K.A., 2010. Comparison of watershed disturbance predictive models for stream benthic macroinvertebrates for three distinct ecoregions in western US. Ecological Indicators 10, 1125–1136.

Whittier, T.R., Stoddard, J.L., Hughes, R.M., Lomnicky, G., 2006. Associations among catchment- and site-scale disturbance indicators and biological assemblages at least- and most-disturbed stream and river sites in the western USA. In: Hughes, R.M., Wang, L., Seelbach, P.W. (Eds.), Landscape influences on Stream Habitat and Biological Assemblages. American Fisheries Society, Symposium 48, pp. 641–664.

Wright, J.F., Moss, D., Armitage, P.D., Furse, M.T., 1984. A preliminary classification of running water sites in Great Britain based on macroinvertebrate species and the prediction of community type using environmental data. Freshwater Biology 14, 221–256.